

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-028945

(43)Date of publication of application : 31.01.1995

(51)Int.Cl.

G06K 9/62

G06K 9/20

G06K 9/32

(21)Application number : 05-030994

(71)Applicant : MATSUSHITA ELECTRIC IND CO
LTD

(22)Date of filing : 19.02.1993

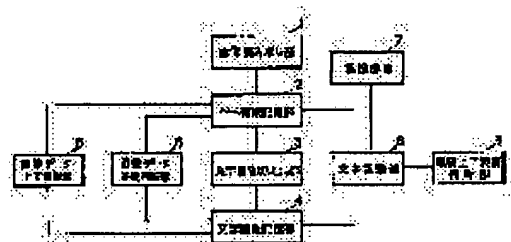
(72)Inventor : ARAMAKI TADASHI

(54) CHARACTER RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To provide a character recognition device which can correctly recognize the characters even when an original is set with its vertical direction reversed or the front and the back of a transparent original is reversed.

CONSTITUTION: A character recognition device consists of an image reading part 1 which reads the image data, a page image storing part 2 which stores the image data in a single page, a character image segmenting part 3 which segments the image data on a single character from the image data on a single page, a character image storing part 4 which stores the image data on a single character, an image data vertically rotating part 5 which vertically rotates the image data by 180°, an image data upper/rear side turning part 6 which reverses the front and the back of the image data, a recognition dictionary 7 which stores the feature information on characters, a character recognition part 8 which obtains the candidate characters and the resemblance degree of the image data, and an original vertical and front/back direction deciding 9 for judging in the vertical and front/back direction of the original.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-28945

(43)公開日 平成7年(1995)1月31日

(51) Int.Cl.⁸

G O B K 9/62
9/20
9/32

識別記号

C 8623-5L
320 N

庁内整理番号

FI

技術表示箇所

審査請求 未請求 請求項の数 1 OL (全 6 頁)

(21)出願番号 特願平5-30994

(22) 出願日 平成5年(1993)2月19日

(71)出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72)發明者 荒 養 正

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

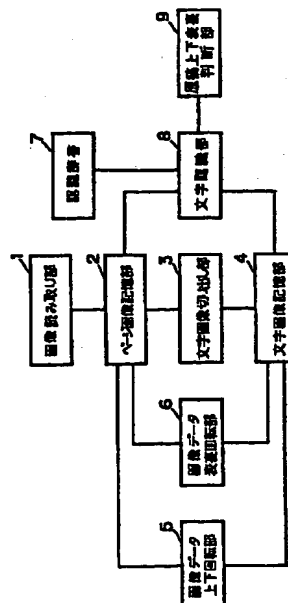
(74) 代理人 弁理士 蔵合 正博

(54) 【発明の名称】 文字認識装置

(57) 【要約】

【目的】 原稿の上下の向きを逆にセットしたり、透明原稿の表裏を逆にセットした場合でも正しく文字認識を行なえるようにする。

【構成】 画像データを読み取るための画像読み取り部1と、1ページの画像データを格納するページ画像記憶部2と、1ページの画像データから1文字の画像データを切り出す文字画像切り出し部3と、1文字の画像データを格納する文字画像記憶部4と、画像データを上下180度回転する画像データ上下回転部5と、画像データの表裏を裏返す画像データ表裏回転部6と、文字の特徴情報を格納した認識辞書7と、画像データの候補文字と類似度を求める文字認識部8と、原稿の上下表裏方向を判断する原稿上下表裏判断部9とを備える。



【特許請求の範囲】

【請求項1】 画像データを読み取るための画像読み取り部と、前記画像読み取り部から入力された1ページの画像データを格納するページ画像記憶部と、前記ページ画像記憶部に格納された1ページの画像データから1文字の画像データを切り出す文字画像切り出し部と、前記文字画像切り出し部から得られた1文字の画像データを格納する文字画像記憶部と、前記ページ画像記憶部に格納された1ページの画像データおよび前記文字画像記憶部に格納された1文字の画像データを上下180度回転する画像データ上下回転部と、前記ページ画像記憶部に格納された1ページの画像データおよび前記文字画像記憶部に格納された1文字の画像データの表裏を裏返す画像データ表裏回転部と、文字の特徴情報を格納した認識辞書と、前記認識辞書に格納された文字の特徴情報を参照して前記文字画像記憶部に格納された画像データの候補文字と類似度を求める文字認識部と、前記文字認識部が求めた候補文字と類似度から原稿の上下表裏方向を判断する原稿上下表裏判断部とを備えた文字認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、入力された画像データから文字データを得る文字認識装置、特にその前処理に関するものである。

【0002】

【従来の技術】 従来、この種の文字認識装置では、新聞、雑誌、論文などの活字、ドット文字および手書き文字等をスキャナなどの光学的読み取り装置で画像データとして読み取り、一文字ずつ分離してJIS等のコード情報に変換して出力していた。

【0003】

【発明が解決しようとする課題】 しかしながら、従来の文字認識装置では、原稿を文字認識装置にセットする時に、原稿の上下の向きが合っていないと正しく文字認識できなかった。また、OHPシートのような透明なシートに書かれた原稿は、表裏と上下の向きが合っていないと正しい認識ができなかった。

【0004】 本発明は、このような従来の問題を解決するものであり、文字認識装置に上下逆にセットされた原稿でも正しく認識することができ、また透明な原稿シートが表裏逆にセットされた場合でも正しく認識することのできる文字認識装置を提供することを目的とする。

【0005】

【課題を解決するための手段】 本発明は、上記目的を達成するために、従来の文字認識装置に、画像データを上下180度回転する画像データ上下回転部と、画像データの表裏を裏返す画像データ表裏回転部と、原稿の上下表裏方向を判断する原稿上下表裏判断部とを加えたものである。

【0006】

【作用】 本発明は、上記構成により、ユーザは文字認識装置にセットする原稿の上下の向きを考える必要がなくなり、また透明な原稿シートの場合は原稿の表裏も考える必要がなくなるので、文字認識作業時のユーザの負担を軽減することができる。

【0007】

【実施例】 以下、本発明の一実施例について、図面を参照しながら説明する。図1は本発明の一実施例における文字認識装置の機能ブロック図である。図1において、1は画像データを読み取るための画像読み取り部、2は画像読み取り部1から入力された1ページの画像データを格納するページ画像記憶部、3はページ画像記憶部2に格納された1ページの画像データから1文字の画像データを切り出す文字画像切り出し部、4は文字画像切り出し部3から得られた1文字の画像データを格納する文字画像記憶部、5はページ画像記憶部2に格納された1ページの画像データおよび文字画像記憶部4に格納された1文字の画像データを上下180度回転する画像データ上下回転部、6はページ画像記憶部2に格納された1ページの画像データおよび文字画像記憶部4に格納された1文字の画像データの表裏を裏返す画像データ表裏回転部、7は文字の特徴情報を格納した認識辞書、8は認識辞書7に格納された文字の特徴情報を参照して文字画像記憶部4に格納された画像データの候補文字と類似度を求める文字認識部、9は文字認識部8が求めた候補文字と類似度から原稿の上下表裏方向を判断する原稿上下表裏判断部である。

【0008】 図2は本発明の一実施例における文字認識装置の回路ブロック図である。図2において、10はランダムアクセスメモリー（以下RAMと略称する。）、11はリードオンリーメモリー（以下、ROMと略称する。）、12は中央処理装置（以下、CPUと略称する。）、13はスキャナである。

【0009】 図1に示した画像読み取り部1は、スキャナ13により、ページ画像記憶部2と文字画像記憶部4は、RAM10により、文字画像切り出し部3と画像データ上下回転部5と画像データ表裏回転部6と原稿上下表裏判断部9は、CPU12がRAM10とデータのやりとりを行ないながらROM11に記憶されたプログラムを実行することにより、認識辞書7は、ROM11により、文字認識部8は、CPU12がROM11やRAM10とデータのやりとりを行ないながらROM11に記憶されたプログラムを実行することにより、それぞれ実現される。

【0010】 以下、本発明の一実施例における文字認識装置について、その動作を図3に示すフローチャートに従って説明する。

【0011】 まずステップAにおいて、ユーザは、画像読み取り部1を用いて画像データを読み取る。入力された1ページの画像データはページ画像記憶部2に格納さ

れる。

【0012】ステップBにおいて、文字画像切り出し部3は、ページ画像記憶部2に格納された1ページの画像データから1文字の画像データを切り出す。得られた1文字のデータは文字画像記憶部4に格納される。

【0013】ステップCにおいて、画像データ上下回転部5は、文字画像記憶部4に記憶された原文画像データから上下回転文字画像データを作る。画像データ表裏回転部6は、文字画像記憶部4に記憶された原文画像データから表裏回転文字画像データを作り、この上下回転文字画像データから上下・表裏回転文字画像データを作る。

【0014】ステップDにおいて、文字認識部8は、原文画像データと上下回転文字画像データと表裏回転文字画像データと上下・表裏回転文字画像データとから各候補文字と類似度を求める。

【0015】ステップEにおいて、原稿上下表裏判断部9は、ステップDで求めた類似度が一番大きな各候補文字が不適文字かどうか判断する。不適文字の場合ステップBに戻る。不適文字でない場合ステップFに進む。

【0016】ステップFにおいて、原稿上下表裏判断部9は、ステップDで求めた類似度が一番大きな画像データを判断する。原文画像データの場合ステップKに進む。上下回転文字画像データの場合ステップGに進む。表裏回転文字画像データの場合ステップIに進む。上下・表裏回転文字画像データの場合ステップH、Jに進む。

【0017】ステップGおよびステップHにおいて、画像データ上下回転部5は、ページ画像記憶部2に格納されている1ページの画像データを上下180度回転する。

【0018】ステップIおよびステップJにおいて、画像データ表裏回転部6は、ページ画像記憶部2に格納されている1ページの画像データの表裏を裏返す。

【0019】ステップKにおいて、文字認識部8は、ページ画像記憶部2に格納されている1ページの画像データを文字認識する。

【0020】以下、本発明の一実施例における文字を認識する様子を図4、図6、図7、図8、図9、図11を用いてさらに詳しく説明する。図4は原稿を上下180度回転する様子を示す。図5は透明な原稿の表裏を裏返す様子を示す。図6は透明な原稿を上下180度回転した後に表裏を裏返す様子を示す。図7はスキャナ13から読み取られたページ画像データを文字画像切り出しして得られた文字画像の例を示す。図8は、図7に示した原文画像を上下180度回転して得られた上下回転文字画像を示す、図9は図7に示した原文画像の表裏を裏返して得られた表裏回転文字画像を示す。図10は図8に示した上下回転文字画像の表裏を裏返して得られた上下・表裏回転文字画像を示す。そして図7、図8、図

9、図10に示した各画像データの文字認識の結果を図11に示す。

【0021】図11から文字認識の結果、類似度が一番高い候補文字は「B」である。ここで得られた候補文字が、

「A」、「M」のような左右がほぼ対象な文字の場合
「C」、「E」のような上下がほぼ対象な文字の場合
「Z」、「N」のようなほぼ点対象な文字の場合
「n」、「u」のようなほぼ点対象な文字がある場合
「b」と「d」と「p」と「q」のようにほぼ対象な文字がある場合

これらの場合は、原稿の向きを判断するには不適切な文字なので、もう一度ページ画像データから文字を切り出す処理に戻る。

【0022】図11から文字認識の結果類似度が一番高い画像データは「上下・表裏回転文字画像」である。このため、スキャナ13にセットされた原稿は上下と表裏が共に反対にセットされていたと判断する。そして、このままページ画像データの文字認識を実行すると正常な結果が得られないので、ページ画像データを上下180度回転した後に表裏を裏返して文字認識を実行する。

【0023】

【発明の効果】本発明は、上記実施例から明らかなように、画像データを上下180度回転する画像データ上下回転部と、画像データの表裏を裏返す画像データ表裏回転部と、原稿の上下表裏方向を判断する原稿上下表裏判断部とを備えているので、ユーザは文字認識装置にセットする原稿の上下の向きを考える必要がなくなり、また透明な原稿シートの場合は原稿の表裏も考える必要がなくなるので、文字認識作業時のユーザの負担を軽減することができる。

【図面の簡単な説明】

【図1】本発明の一実施例における文字認識装置の機能ブロック図。

【図2】本発明の一実施例における文字認識装置の回路ブロック図。

【図3】本発明の一実施例における文字認識装置のフローチャート。

【図4】本発明の一実施例における原稿の上下を回転させた状態の模式図。

—【図5】本発明の一実施例における原稿の表裏を回転させた状態の模式図。

【図6】本発明の一実施例における原稿の上下および表裏を回転させた状態の模式図。

【図7】本発明の一実施例における原文字画像図。

【図8】本発明の一実施例における上下回転文字画像図。

【図9】本発明の一実施例における表裏回転文字画像図。

【図10】本発明の一実施例における上下および表裏回

転文字画像図。

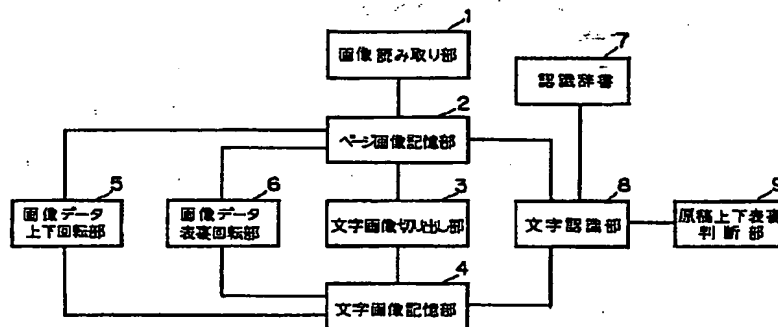
【図11】本発明の一実施例における候補文字と類似度の関係を示す一覧図。

【符号の説明】

- 1 画像読み取り部
- 2 ページ画像記憶部
- 3 文字画像切り出し部
- 4 文字画像記憶部
- 5 画像データ上下回転部

- 6 画像データ表裏回転部
- 7 認識辞書
- 8 文字認識部
- 9 原稿上下表裏判定部
- 10 RAM
- 11 ROM
- 12 CPU
- 13 スキャナ

【図1】



【図7】

B

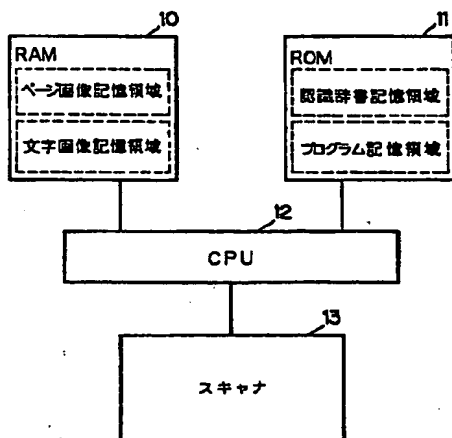
【図8】

B

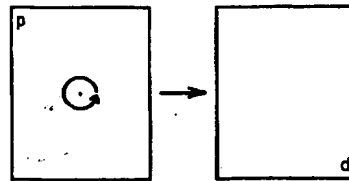
【図9】

B

【図2】



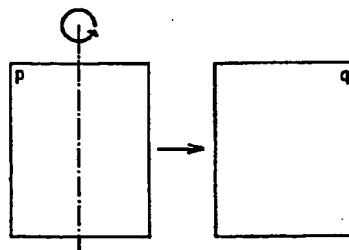
【図4】



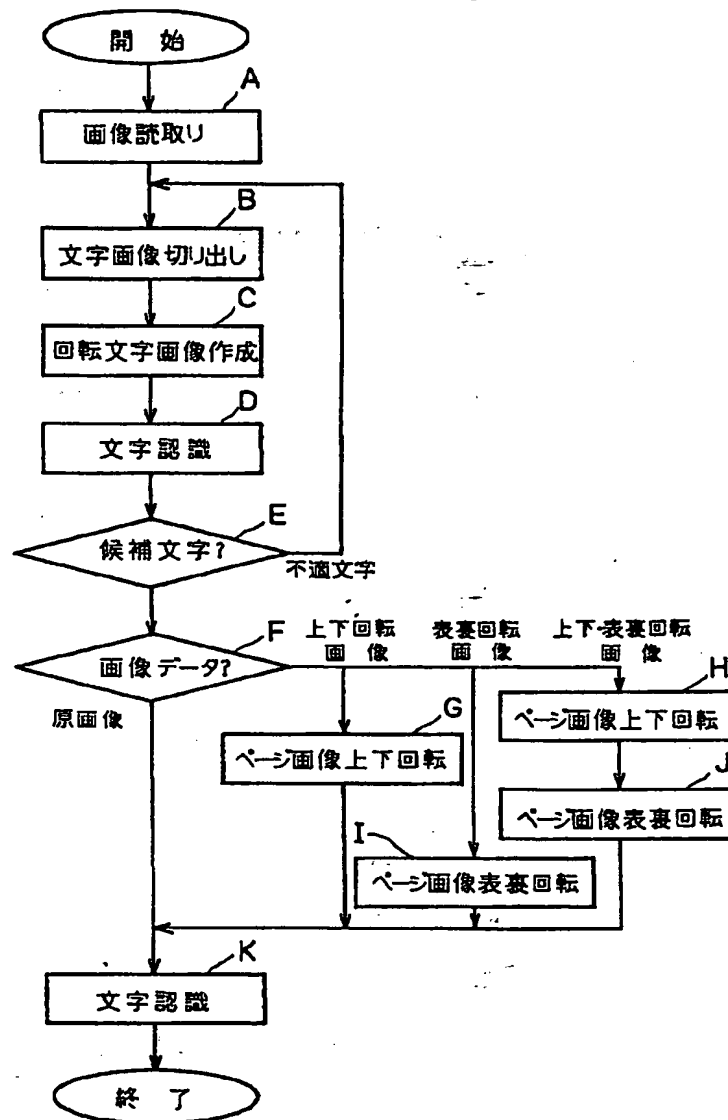
【図10】

B

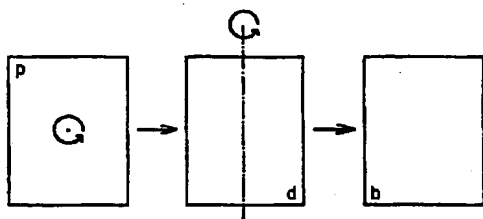
【図5】



【図3】



【図6】



【図11】

文字画像	候補文字	類似度
原文字画像	B	0.91
上下回転文字画像	8	0.77
表裏回転文字画像	8	0.71
上下・表裏回転文字画像	B	0.96

Translation of laid-open Japanese patent

Japanese Patent Application: Publication No: H07-28945 (A)
Publication date: Jan. 31, 1995
Filing No: H05-30994
Filing date: Feb. 19, 1993

Applicant: Matsushita Electric Industrial Co. Ltd

Inventor: Tadashi ARAMAKI

Patent Attorney: Masahiro KURAAI

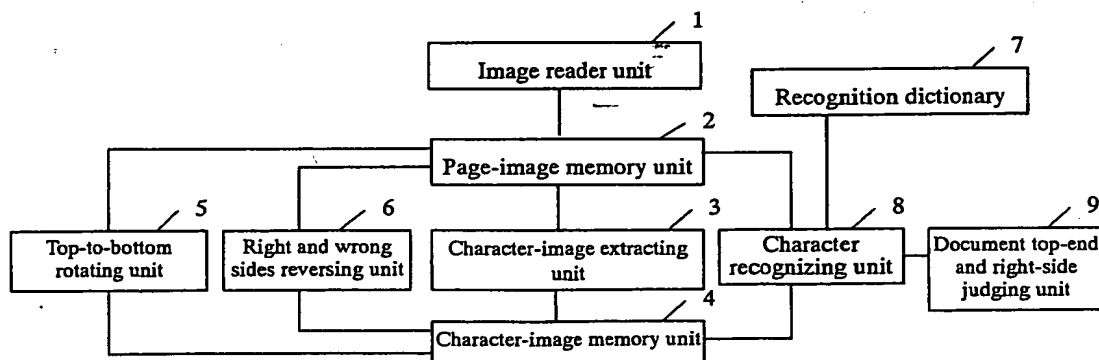
[Title] Character Recognition Apparatus

[Abstract]

[Objective] To enable an apparatus recognizing characters on a document page even when the sheet of the document page being placed in the wrong top-to-bottom direction or in a wrong side up manner.

[Constitution] The character recognition apparatus is comprised of:

- an image reader unit 1,
- a page-image memory unit 2 for storing an image-data set covering the image of an entire page,
- a character-image extracting unit 3 for extracting from the image-data set covering the image of the entire page, a portion correspondent with an individual character,
- a character-image memory unit 4 for storing the image-data portion correspondent with the individual character,
- a top-to-bottom rotating unit 5 for rotating 180 degrees the image of the image-data set or image-data portion so that the top and bottom of the image is inverted,
- a right and wrong side reversing unit 6 for reversing sides of the page image of the image-data set,
- a recognition dictionary 7 containing information about features associated with respective characters,
- a character recognizing unit 8 for obtaining character candidates in relation to the image data set together with respectively associated similarity degrees,
- a document top-end and right-side judging unit 9 for judging top-and-bottom ends and right-and-wrong sides of a document.



[Scope of claims]

Claim 1. A character recognition apparatus comprising:

- an image reader unit for acquiring image data,
- a page-image memory unit for storing an image-data set acquired by the image reader unit and covering the image of an entire page,
- a character-image extracting unit for extracting from the image-data set that covers the image of the entire page, a portion correspondent with an individual character,
- a character-image memory unit for storing the image-data portion, which is extracted by the character-image extracting unit and correspondent with the individual character,
- a top-to-bottom rotating unit for rotating 180 degrees the entire-page image of the image-data set stored in the page-image memory unit or the individual-character image of the image-data portion stored in the character-image memory unit so that the top and bottom of either image is inverted,
- a right and wrong side reversing unit for reversing sides of the entire-page image of the image-data set stored in the page-image memory unit or sides of the individual-character image of the image-data portion stored in the character-image memory unit,
- a recognition dictionary containing information about features associated with respective characters,
- a character recognizing unit for obtaining character candidates correspondent with the image-data portion stored in the character-image memory unit together with similarity degrees associated with the character candidates respectively by making reference to the information, which is about the features associated with respective characters and contained in the recognition dictionary, and
- a document top-end and right-side judging unit for judging top-and-bottom ends and right-and-wrong sides of a document based on the character candidates and associated similarity degrees obtained by the character recognizing unit.

[Detailed explanation of invention]

[0001]

[Field of industrial application] The invention is concerned with a character recognition apparatus for converting an acquired set of image data into a set of character data, and in particular, with a pre-process performed by the character recognition apparatus.

[0002]

[Prior art technology] Character recognition apparatuses of this kind and associated with prior art, each acquires printed, dot-matrix or handwritten characters contained in a newspaper, magazine, report or the like by means of an optical image-reader unit such as a scanner as a set of image data and outputs them in a set of coded data, wherein the image data is divided into portions respectively representing individual characters and converted into respectively corresponding character codes in compliant with a standardized rule such as JIS (Japanese Industrial Standards).

[0003]

[Problem to be solved] A character recognition apparatus according to prior art is unable to perform a character recognition process correctly if a subject document page is placed on it wrongly with respect to the top-to-bottom end direction. When in a situation in which the subject document page is of a transparent sheet such as an OHP sheet, it is necessary to place it correctly with respect to both the top-to-bottom end and right-side facing directions for obtaining a successful recognition result.

[0004] Present invention resolves this problem associated with prior art. The present invention makes it the objective to provide a character recognition apparatus that can perform a character recognition process correctly either in a case in which a subject document page is placed wrongly in terms of the top-to-bottom end direction or in a case in which a transparent sheet constituting a subject document is placed in a wrongly faced manner on the character recognition apparatus.

[0005]

[Means for solving problem] The character recognition apparatus according to the present invention, in order for achieving the above-mentioned objective, comprises, further to those constituting a prior art character recognition apparatus:

a top-to-bottom rotating unit for rotating 180 degrees the image of an image-data set so that the top and bottom ends of the image is inverted,

a right and wrong side reversing unit for reversing sides of the image of an image-data set, and

a document top-end and right-side judging unit for judging top-and-bottom ends and right-and-wrong sides of a document.

[0006]

[Benefit] Work load to a user placing subject documents to the character recognition apparatus reduces according to the present invention, because it is unnecessary to be specific about the top-to-bottom end direction when placing a subject document sheet, and further, if the subject document sheet is constituted by a transparent sheet, it is unnecessary to be specific about the right-or-wrong-side facing direction of the sheet, either.

[0007]

[Embodiment] An embodiment of the present invention is described below with references made to drawings. Fig.1 is a diagram of function blocks associated with a character recognition apparatus representing an embodiment of the present invention. In Fig.1, 1 indicates the image reader unit, which is for acquiring an image data set, 2 indicates the page-image memory unit, which is for storing the image data-set acquired by the image reader unit 1 and covering the image of an entire page, 3 indicates the character-image extracting unit, which is for extracting the image-data portion that is correspondent with an individual character from the image-data set that is acquired by the image reader unit 1 and covers the image of an entire page, 4 indicates the character-image data memory unit, which is for storing the image-data portion that is extracted by the character-image extracting unit 3 and is correspondent with an

individual character, 5 indicates the top-to-bottom rotating unit, which is for rotating 180 degrees the entire-page image of the image-data set stored in the page-image memory unit 2 or the individual-character image of the image-data portion stored in the character-image memory unit 4, 6 indicates the right and wrong side reversing unit, which reverses sides of the entire-page image of the image-data set stored in the page-image memory unit 2 or sides of the individual-character image of the image-data portion stored in the character-image memory unit 4, 7 indicates the recognition dictionary, which contains information about features associated with respective characters, 8 indicates the character recognizing unit, which is for obtaining character candidates correspondent with the image-data portion stored in the character-image memory unit 4 together with similarity degrees associated respectively with the character candidates by making reference to the information, which is about the features associated with respective characters and contained in the recognition dictionary 7, and 9 indicates the document top-end and right-side judging unit for judging top-and-bottom ends and right-and-wrong sides of a document based on the character candidates and associated similarity degrees obtained by the character recognizing unit 8.

[0008] Fig.2 is a diagram of circuit blocks constituting the character recognition apparatus of a present invention embodiment. In Fig.2, 10 indicates a random-access-memory (hereafter called RAM), 11 indicates a read-only-memory (hereafter called ROM), 12 indicates a central processing unit (hereafter called CPU) and 13 indicates a scanner.

[0009] Circuit blocks are employed to implement respective units shown in Fig.1 as follows. The image reader unit 1 is implemented by a scanner 13, the page-image memory unit 2 and character-image memory unit 4 are implemented by a RAM 10, the character-image extracting unit 3, the top-to-bottom rotating unit 5, the right and wrong sides reversing unit 6 and the document top-end and right-side judging unit 9 are implemented by operation of a CPU 12 that executes a computer-program stored in the ROM 11 and exchanges data with the RAM 10, the recognition dictionary 7 is implemented by the ROM 11, and the character recognizing unit 8 is implemented by the CPU 12 that executes a computer-program stored in the ROM 11 and exchanges data with the ROM 11 and RAM 10.

[0010] The operation of the character recognition apparatus of a present invention embodiment is as described below in a manner associated with the flowchart shown in Fig.3.

[0011] First, at step A, a user obtains an image-data set using an image reader unit 1. The image-data set thus obtained and covering the image of an entire page is stored in a page-image memory unit 2.

[0012] At step B, a character-image extracting unit 3 extracts an image-data portion correspondent with an individual character from the image-data set stored in the page-image memory unit 2 and covering the image of an entire page. The data correspondent with the individual character is stored in a character-image memory unit 4.

[0013] At step C, a top-to-bottom rotating unit 5 generates a rotated-character image-data portion from the original image-data portion stored in the character-image memory unit 4. A right and wrong sides

reversing unit 6 generates a side-reversed-character image-data portion from the original image-data portion stored in the character-image memory unit 4 and separately a top-end right-side inverted image-data portion from the rotated-character image-data portion.

[0014] At step D, a character-recognizing unit 8 obtains character candidates and associated similarity degrees for the original image-data portion, rotated-character image-data portion, side-reversed-character image-data portion and top-end right-side inverted image-data portion, respectively.

[0015] At step E, the document top-end and right-side judging unit 9 judges whether the character of which the associated similarity degree obtained in step D is the highest is an unacceptable character. If it is judged to be an unacceptable character, the operation goes back to step B. Otherwise, it proceeds to step F.

[0016] At step F, the document top-end and right-side judging unit 9 evaluates the image-data portion representing the character of which the associated similarity degree obtained in step D is the highest. If the image-data portion is the original image-data portion, the operation moves to step K. If it is the rotated-character image-data portion, the operation moves to step G. If it is the side-reversed-character image-data portion, the operation moves to step I. And if it is the top-end right-side inverted image-data portion, the operation moves to steps H and J.

[0017] At steps G and H, the top-to-bottom rotating unit 5 rotates 180 degrees the entire page image defined by the image-data set, which is stored in the page-image memory unit 2 so that the top and bottom ends of the associate page image are interchanged.

[0018] At steps I and J, the right and wrong sides reversing unit 6 reverses the sides of the entire page image that is defined by the image-data set stored in the page-image memory unit 2.

[0019] At step K, the character recognizing unit 8 performs character recognizing processes with respect to the image-data set stored in the page-image memory unit 2 and covering an entire page.

[0020] Here, the manner in which characters are recognized in a present invention embodiment is provided more specifically using drawings provided in Figs. 4, 6, 7, 8, 9 and 11. The drawing in Fig.4 is an illustration to show how a document image is rotated 180 degrees to interchange its top and bottom ends. The drawing in Fig.5 is an illustration to show how a transparent document image is reversed between its right and wrong sides. The drawing in Fig.6 illustrates that a transparent document image is first rotated 180 degrees to interchange its top and bottom ends and then is reversed between its right and wrong sides. The drawing in Fig.7 is an example of a character image defined by an image-data portion resulting from a character image extraction process performed over an image-data set, which is obtained by a scanner 13 and covers an entire page. The drawing in Fig.8 is the top-to-bottom rotated character image resulting from 180-degree rotation of the original image shown in Fig.7. The drawing in Fig.9 is the side reversed character image resulting from right and wrong side reversal of the original image shown in Fig.7. The drawing in Fig.10 is the top-and-bottom ends right-and-wrong sides inverted character image resulting from right and wrong side reversal of the top-to-bottom rotated character image

shown in Fig.8. Lastly, Fig.11 shows the result of a character recognition process performed over image-data sets respectively representing images shown in Figs.7, 8, 9 and 10.

[0021] The character recognition performed basing on data contained in the table shown in Fig.11 identifies character "B" as the one associated with the highest similarity degree among character candidates. In a case in which the character candidate identified as the one associated with the highest similarity degree is any of the following characters, the operation returns to a step in which a character image is extracted from the image-data set covering an entire page, as any of these characters are unacceptable for use in judging the placement direction of a document.

Characters virtually symmetric about the vertical line at the center. Examples are "A" and "M".

Characters virtually symmetric about the horizontal line at the center. Examples are "C" and "E".

Characters virtually symmetric around the center point. Examples are "Z" and "N".

Characters of which the rotated variations resemble to other characters. An example is "n" of which the rotated variation resembles "u".

Characters of which the mirror images constitute other characters. Examples are character pairs of "b" and "d", and "p" and "q".

[0022] According to the table shown in Fig.11, the character recognition, here, has yielded the result that "the top-end right-side inverted character image" has the highest similarity degree. Based on this result the apparatus judges that the document has been placed on the scanner 13 wrongly both in the top-and-bottom direction and in the side selection. As it is impossible to produce a correct character recognition result by performing character recognition over the image-data set of which the correspondent page-image is in this state, the image-data set is modified into one to represent the page-image which is 180-degree rotated and then turned over with respect to the document sides before commencing the character recognition.

[0023]

[Effect of invention]

According to the present invention, as it becomes clear from the foregoing, the character recognition apparatus is further provided with a top-to-bottom rotating unit for rotating 180 degrees an image associated with an image-data set so that the top and bottom ends of the image is converted, a right and wrong sides reversing unit for reversing sides of an image associated with an image-data set and a document top-end and right-side judging unit for judging top-and-bottom ends and right-and wrong sides of a document. Consequently it becomes unnecessary for a user to be specific about the top-to-bottom direction in which a subject document sheet is placed, and further about the right-or-wrong-side facing direction when if the subject document sheet is constituted by a transparent sheet.

Work load to a user placing subject documents to the character recognition apparatus reduces in this way.

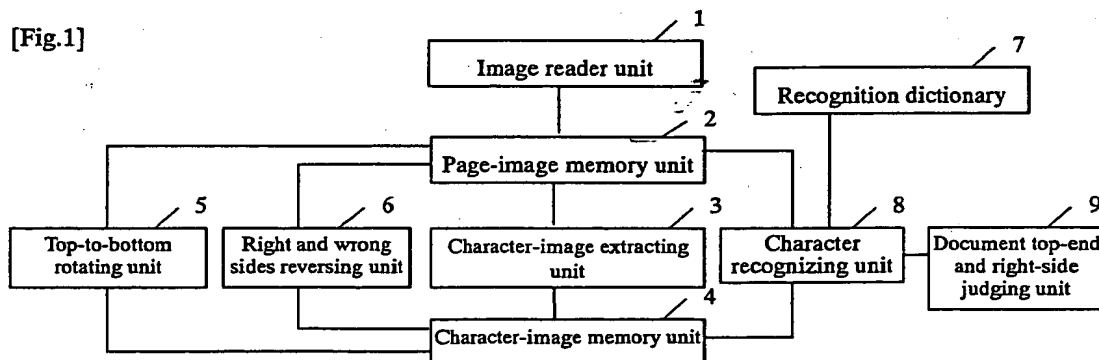
[Brief explanation to drawings]

- [Fig.1] Functional-block diagram of a present invention embodiment, a character recognition apparatus
- [Fig.2] Circuit-block diagram of a present invention embodiment, a character recognition apparatus
- [Fig.3] Flowchart of a present invention embodiment, a character recognition apparatus
- [Fig.4] Schematic illustration of a document being rotated to have its top and bottom ends interchanged according to the manner associated with a present invention embodiment
- [Fig.5] Schematic illustration of a document being turned over to have its right and wrong sides interchanged according to the manner associated with a present invention embodiment
- [Fig.6] Schematic illustration of a document being rotated to have its top and bottom ends interchanged and further turned over to have its right and wrong sides interchanged according to the manner associated with a present invention embodiment
- [Fig.7] Image of an original character associated with a present invention embodiment
- [Fig.8] Character-image of the original character of which the top and bottom ends are interchanged in association with a present invention embodiment
- [Fig.9] Character-image of the original character of which the right and wrong sides are interchanged in association with a present invention embodiment
- [Fig.10] Character-image of the original character of which the top and bottom ends are interchanged and further its right and wrong sides are interchanged in association with a present invention embodiment
- [Fig.11] Tabular description of the relation between character candidates and associated similarity degrees according to the manner associated with a present invention embodiment

[Explanation of indexes]

- | | |
|---|---|
| 1: Image reader unit | 8: Character recognizing unit |
| 2: Page-image memory unit | 9: Document top-end and right-side judging unit |
| 3: Character-image extracting unit | 10: RAM |
| 4: Character-image memory unit | 11: ROM |
| 5: Top-to-bottom rotating unit | 12: CPU |
| 6: Right and wrong sides reversing unit | 13: Scanner |
| 7: Recognition dictionary | |

[Fig.1]



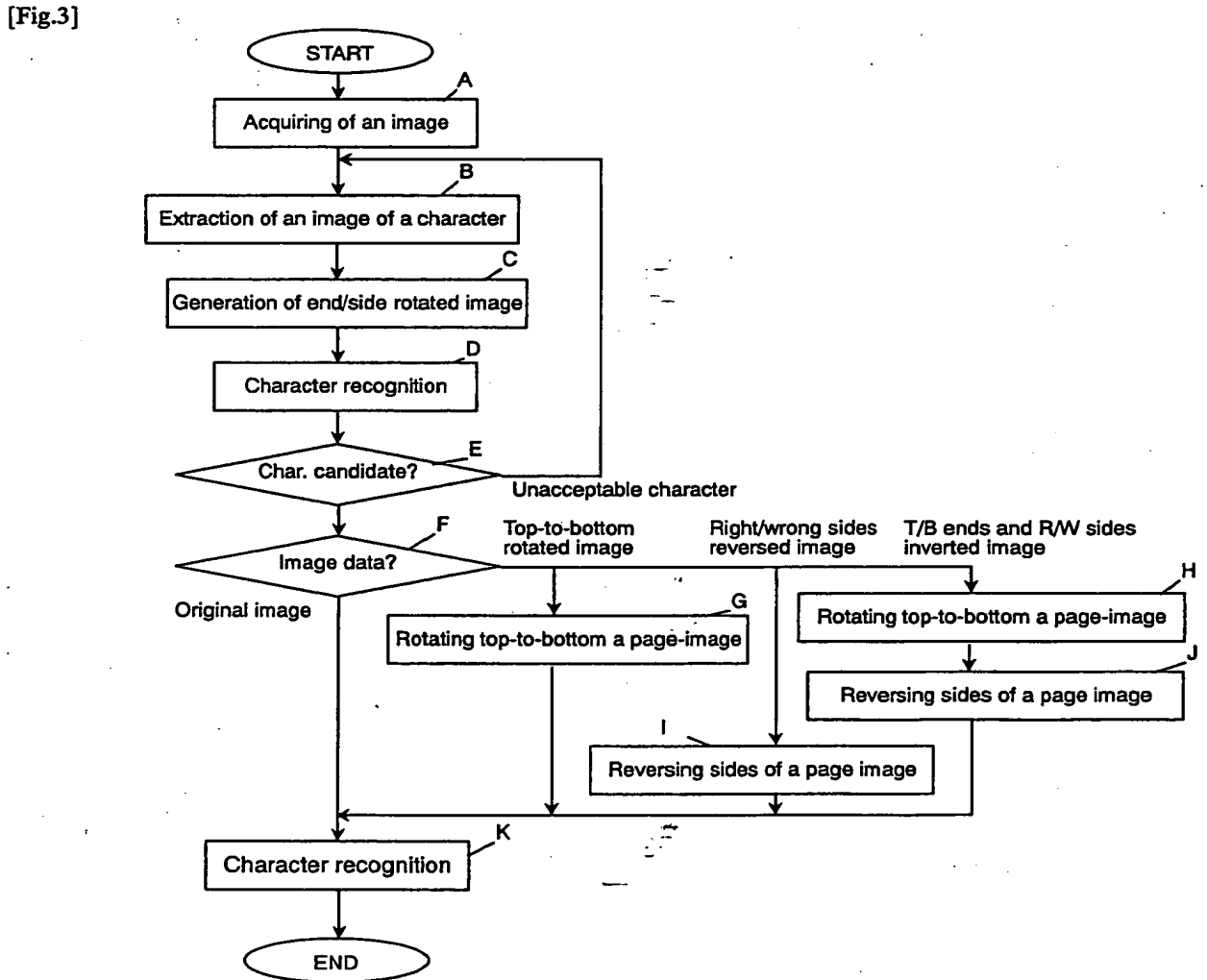
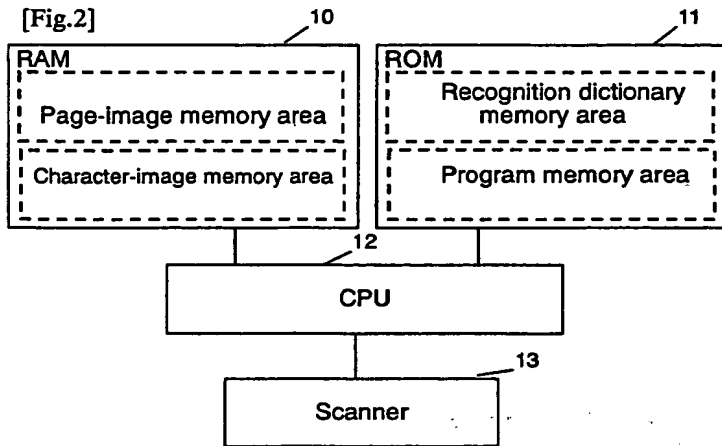


Fig.4

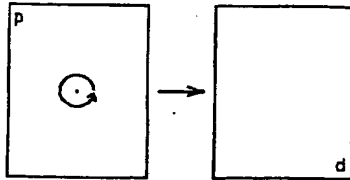


Fig.5

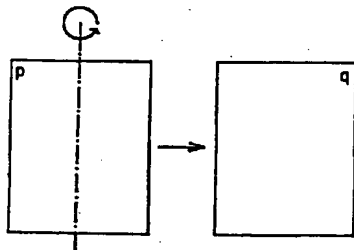


Fig.7

B

Fig.8

B

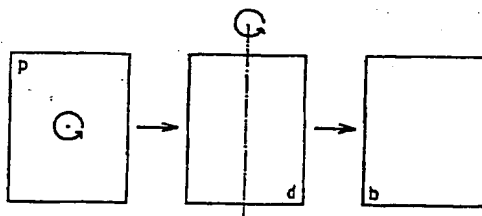
Fig.9

B

Fig.10

B

Fig.6



[Fig.11]

Character image	Character candidate	Similarity degree
Original image of the character	B	0.91
Top-to-bottom rotated image of the character	8	0.77
Right/wrong sides reversed image of the character	8	0.71
Top/bottom ends and right/wrong sides converted image of the character	B	0.96